**Assignment 04 – HMM POS tagging**

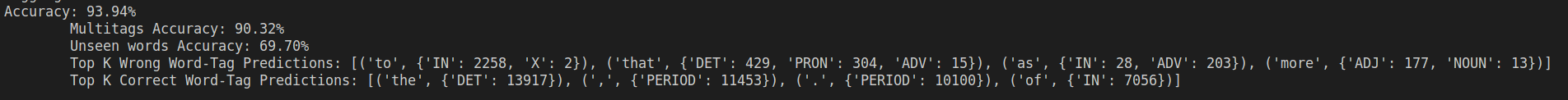
***Chentai Yuan, Qianzhong Chen, Hao Ding***

***ECE448, Group 07***

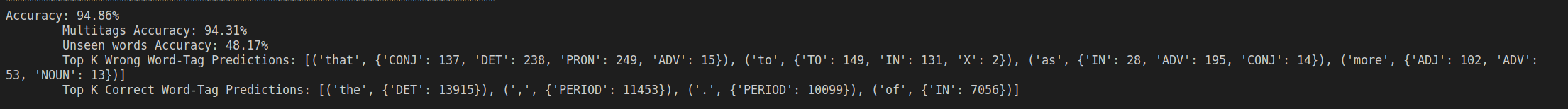
***TA: Shuting Tao, Hanrong Zhang***

***May 15, 2023***

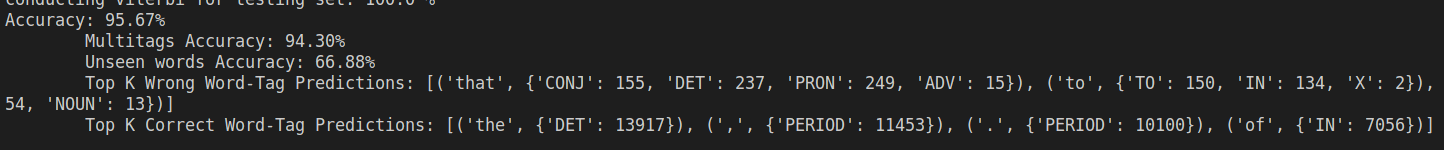
**Section Ⅰ: Baseline**



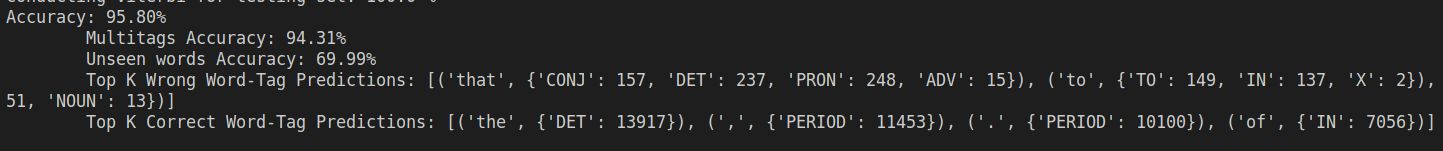
**Section Ⅱ: Viterbi\_1**



**Section Ⅲ: Viterbi\_2**

****

**Section Ⅳ: Viterbi\_ec**



In this section, we built a suffix & Prefix set includes [X\_ING, X\_ED, X\_LY, UN\_X, IN\_X]. From hapax word set, we count the probability of P(tag | fix), i.e., if a prefix or suffix is detected from an unknown word, the probability that the word has a certain tag. As we chose the prefix and suffix that are widely used, our new algorithm has achieved satisfied prediction performance.

**ACKNOWLEGEMENTS**

Qianzhong Chen wrote his own version of baseline.py, Viterbi\_1.py, Viterbi\_2.py, Viterbi\_ec.py, and amend the provided Viterbi\_1.py file.

Chentai Yuan try to debug with the Viterbi\_1.py and Viterbi\_2.py.

Hao Ding assist on writing the Viterbi\_1.py, Viterbi\_2.py.